Johnson, O., & Guha, S. (2017). A de Bruijn identity for discrete random variables. In *2017 IEEE International Symposium on Information Theory (ISIT 2017)* (pp. 898-902). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/ISIT.2017.8006658

A de Bruijn identity for discrete random variables

Oliver Johnson
School of Mathematics
University of Bristol
Bristol, BS8 1TW, UK
Email: maotj@bristol.ac.uk

Saikat Guha
Quantum Information Processing group
Raytheon BBN Technologies
Cambridge, MA 02138, USA
Email: saikat.guha@raytheon.com

***Abstract*—We discuss properties of the "beamsplitter addition" operation, which provides a non-standard scaled convolution of random variables supported on the non-negative integers. We give a simple expression for the action of beamsplitter addition using generating functions. We use this to give a self-contained and purely classical proof of a heat equation and de Bruijn identity, satisfied when one of the variables is geometric.**

I. Introduction and notation

Stam [1] showed that addition of independent continuous random variables satisfies the de Bruijn identity [1, Eq. (2.12)], in that the derivative of entropy under the addition of a normal is Fisher Information. This identity underpins many analyses of entropy under addition, including Stam's proof of Shannon's Entropy Power Inequality (EPI) [2], Barron's information theoretic Central Limit Theorem [3], and Madiman and Barron's proof of monotonicity of entropy under addition of independent identically-distributed (i.i.d.) random variables [4]. These results have the Gaussian distribution at their heart, relating to the Gaussian maximum entropy property and closure of the Gaussian family under addition ("Gaussian + Gaussian = Gaussian"). The de Bruijn identity follows since the densities in question satisfy the heat equation [1, Eq. (5.1)].

There have been many attempts to develop a corresponding theory for discrete random variables, often focussing on the Poisson family which is closed under standard integer addition ("Poisson + Poisson = Poisson"). Results in this context include Poisson limit theorems [5], [6], maximum entropy property [7] and monotonicity result [8]. However, [7] and [8] rely on the assumption of ultra-log-concavity (ULC), meaning that they are more restrictive than their Gaussian counterparts.

In this paper we prove new properties of what we refer to as the 'beamsplitter addition' $\boxplus_\eta$ (see Definition I.1) of random variables supported on the non-negative integers $\mathbb{Z}_+$. By design, the geometric family is closed under the action of $\boxplus_\eta$ ("Geometric $\boxplus_\eta$ Geometric = Geometric"). Geometric is more natural than the Poisson since no auxiliary assumptions such as ULC are required to prove maximum entropy.

The beamsplitter addition is motivated by how an optical beamsplitter of transmissivity $\eta \in [0,1]$ "adds" the photon-number distributions of two classical mixtures of number states. It underlies the conjectural Entropy Photon Number Inequality (EPnI) [9], [10], which plays a role analogous to Shannon's EPI in understanding the capacity of Gaussian bosonic channels. The paper [11] includes a more detailed history of the beamsplitter addition $\boxplus_\eta$. Our key aim however is to give a self-contained presentation of the beamsplitter addition $\boxplus_\eta$ as a way of combining random variables supported on $\mathbb{Z}_+$, accessible to a purely classical audience. Although some of our results may be known to the quantum information community, we hope this paper will stimulate future work by probabilists and classical information theorists on open problems—in particular, a proof of the conjectured EPI and entropic monotonicity under the beamsplitter addition [12].

For brevity, we assume henceforth that random variables are supported on $\mathbb{Z}_+$ and relevant pairs of random variables are independent. We first define $\boxplus_\eta$ in the notation of [12]:

Definition I.1. *Given a random variable X, define its continuous counterpart $\mathbf{X}_c$ (a circularly symmetric random variable supported on the complex plane $\mathbb{C}$), using a map $\mathcal{T}$ with $\mathbf{X}_c = \mathcal{T}(X)$ and $X = \mathcal{T}^{-1}(\mathbf{X}_c)$ with actions on mass functions and densities given by [12, Eq. (14),(15)]:*

$$p_{\mathbf{X}_c}(\mathbf{r}) = \sum_{n=0}^{\infty} p_X[n] \frac{e^{-|\mathbf{r}|^2}|\mathbf{r}|^{2n}}{n!\pi}, \tag{1}$$

$$p_X[n] = \frac{1}{\pi}\int_{\mathbb{C}} p_{\mathbf{X}_c}(\mathbf{r})\mathcal{L}_n\left(|\mathbf{s}|^2\right)\exp(\mathbf{r}\mathbf{s}^* - \mathbf{r}^*\mathbf{s})d\mathbf{r}d\mathbf{s}, \tag{2}$$

where $\mathcal{L}_n$ denotes the n^{th} Laguerre polynomial. As in [12], for $0 \leq \eta \leq 1$, we define the beamsplitter addition operation $\boxplus_\eta$ acting on independent X, Y supported on $\mathbb{Z}_+$ by

$$X \boxplus_\eta Y = \mathcal{T}^{-1}\left(\sqrt{\eta}\,\mathcal{T}(X) + \sqrt{1-\eta}\,\mathcal{T}(Y)\right), \tag{3}$$

where '+' on the RHS of (3) denotes standard addition in $\mathbb{C}$.

The key contributions of the paper are as follows. In Section II we define two types of generating functions, for X and $\mathbf{X}_c$, and prove a new relation between them in Theorem II.4. In Section III, we prove Theorem III.1, which shows that the generating function of $X \boxplus_\eta Y$ is a product of generating functions. In Section IV, we show that Theorem III.1 implies a heat equation (Theorem IV.1), which in turn gives a de Bruijn identity (Theorem IV.3). In Section V, we state and prove a new bound on relative entropy under the action of $\mathcal{T}$. In Section VI we discuss future work. The remainder of the paper contains proofs of its main results.

II. Relation between generating functions

We consider two different kinds of generating functions, exponential and ordinary, recalling that they are related by the

Laplace transform (see Lemma VII.1). We write $\mathbb{E}\,(X)_{(m)} := \mathbb{E}X(X-1)\dots(X-m+1) = \mathbb{E}X!/(X-m)!$ for the falling moment of a random variable on $\mathbb{Z}_+$.

Definition II.1.

1) *Given random variable X with probability mass function (p.m.f.) $p_X[m]$, $m \in \mathbb{Z}_+$, consider the sequence $\mathbb{E}\,(X)_{(m)}/m!$ and write:*
   a) *The ordinary generating function*
   $$\psi_X(t) := \sum_{m=0}^{\infty} t^m \left(\frac{\mathbb{E}\,(X)_{(m)}}{m!} \right), \qquad (4)$$
   b) *The exponential generating function*
   $$\widetilde{\psi}_X(t) := \sum_{m=0}^{\infty} \frac{t^m}{m!} \left(\frac{\mathbb{E}\,(X)_{(m)}}{m!} \right). \qquad (5)$$
2) *For circularly symmetric $\mathbf{X}_c$ supported on $\mathbb{C}$ with density $p_{\mathbf{X}_c}$, consider the sequence $\frac{\mathbb{E}|\mathbf{X}_c|^{2m}}{m!}$ and write:*
   a) *The ordinary generating function*
   $$\phi_{\mathbf{X}_c}(t) \;:=\; \sum_{m=0}^{\infty} t^m \left(\frac{\mathbb{E}|\mathbf{X}_c|^{2m}}{m!} \right), \qquad (6)$$
   b) *The exponential generating function*
   $$\widetilde{\phi}_{\mathbf{X}_c}(t) := \sum_{m=0}^{\infty} \frac{t^m}{m!} \left(\frac{\mathbb{E}|\mathbf{X}_c|^{2m}}{m!} \right). \qquad (7)$$

Note that although (4) and (5) are defined as formal sums, in practice we focus on $t \le 0$. We first make the following claim, proved in Section VII:

Lemma II.2. *For any random variable X supported on $\mathbb{Z}_+$:*
1) $\widetilde{\psi}_X(t) = \sum_{n=0}^{\infty} p_X[n] \mathcal{L}_n(-t)$,
2) $p_X[m] = \int_0^{\infty} \exp(-s) \widetilde{\psi}_X(-s) \mathcal{L}_m(s) ds$.

Example II.3. *For X geometric on $\{0,1,\dots\}$ with mean λ, $\mathbb{E}\,(X)_{(m)} = m!\lambda^m$, so $\psi_X(t) = 1/(1-\lambda t)$ and $\widetilde{\psi}_X(t) = \exp(\lambda t)$. Further $\mathbf{X}_c$ is circularly symmetric Gaussian with covariance matrix $(1+\lambda)\mathbf{I}_2/2$, where we write $\mathbf{I}_d$ for the d-dimensional identity matrix. Hence, $\mathbb{E}|\mathbf{X}_c|^{2m} = m!(1+\lambda)^m$, so $\phi_{\mathbf{X}_c}(t) = 1/(1-(1+\lambda)t)$ and $\widetilde{\phi}_{\mathbf{X}_c}(t) = \exp(t(1+\lambda))$.*

Example II.3 illustrates the following result, in that X and $\mathbf{X}_c$ are linked at the level of their generating functions:

Theorem II.4. *For X and $\mathbf{X}_c$ linked by the transforms* (1) *and* (2) *of [12], we can write*
1)
$$\widetilde{\psi}_X(t) = \exp(-t) \widetilde{\phi}_{\mathbf{X}_c}(t). \qquad (8)$$
2)
$$\psi_X(t) = \frac{1}{1+t} \phi_{\mathbf{X}_c} \left(\frac{t}{1+t} \right), \qquad (9)$$

Proof: See Section VII. ■

This result relates the moments of X and $\mathbf{X}_c$. For brevity, from now on we write λ_W for the mean of any random variable W. Then, for example:

Corollary II.5. *The real and imaginary parts of $\mathbf{X}_c = (X_1, X_2)$ have covariance matrix $(1+\lambda_X)\mathbf{I}_2/2$.*

Proof: We simply differentiate (9) with respect to t and set $t = 0$ to obtain:
$$\lambda_X = \psi'_X(0) = \phi'_{\mathbf{X}_c}(0) - \phi_{\mathbf{X}_c}(0) = \mathbb{E}|\mathbf{X}_c|^2 - 1.$$
Since $\mathbf{X}_c$ is circularly symmetric, it is proper (see [14]), and we know $\mathbb{E}\mathbf{X}_c = 0$. Further, this means that (X_1, X_2) has a covariance matrix which is a multiple of the identity. We deduce that the diagonal entries must equal $(\lambda_X + 1)/2$. ■

III. Generating functions and beamsplitter addition

We now state the relationship between the generating functions of X, Y and $Z = X \boxplus_\eta Y$, proved in Section VIII:

Theorem III.1. *Given independent random variables X and Y supported on $\mathbb{Z}_+$, the $Z := X \boxplus_\eta Y$ has generating functions $\widetilde{\psi}_Z$ and ψ_Z satisfying:*
1)
$$\widetilde{\psi}_Z(t) = \widetilde{\psi}_X(\eta t) \widetilde{\psi}_Y((1-\eta)t). \qquad (10)$$
2)
$$\frac{1}{s-1} \psi_Z \left(\frac{1}{s-1} \right) = \left[L \left(M_X^{(\eta)} \times M_Y^{(1-\eta)} \right) \right](s), \qquad (11)$$
where we define $M_X^{(\eta)}$ and $M_Y^{(1-\eta)}$ via the inverse Laplace transform $\left[L^{-1} \cdot\right]$ using the fact that:
$$\left[L M_{\mathbf{X}_c}^{(\eta)} \right](s) \;=\; \frac{1}{s-\eta} \psi_X \left(\frac{\eta}{s-\eta} \right), \qquad (12)$$
$$\left[L M_{\mathbf{Y}_c}^{(1-\eta)} \right](s) \;=\; \frac{1}{s-(1-\eta)} \psi_Y \left(\frac{1-\eta}{s-(1-\eta)} \right) \qquad (13)$$

Direct calculation of the derivative of (10), as in Corollary II.5, allows us to deduce that
$$\lambda_Z = \eta \lambda_X + (1-\eta) \lambda_Y. \qquad (14)$$

Example III.2. *If X is geometric with mean λ_X and Y is geometric with mean λ_Y, using the expressions from Example II.3 and* (14) *then*
1) *The RHS of* (10) *becomes*
$$\exp(\eta \lambda_X t) \exp((1-\eta)\lambda_Y t) = \exp(\lambda_Z t),$$
so that $Z = X \boxplus_\eta Y$ is geometric with mean λ_Z.
2) *The RHS of* (12) *is $1/(s - \eta(1+\lambda_X))$, so the inverse Laplace transform gives that $M_{\mathbf{X}_c}^{(\eta)}(t) = \exp(\eta(1+\lambda_X)t)$, with $M_{\mathbf{Y}_c}^{(\eta)}(t) = \exp((1-\eta)(1+\lambda_Y)t)$. As we would expect, this means that $\left(M_X^{(\eta)} \times M_Y^{(1-\eta)}\right) = \exp((1+\lambda_Z)t)$ This allows us to deduce that*
$$\frac{1}{s-1} \psi_Z \left(\frac{1}{s-1} \right) = \frac{1}{s-(1+\lambda_Z)},$$
and changing variables via $u = 1/(s-1)$ we deduce that $\psi_Z(u) = 1/(1-\lambda_Z u)$ as we would hope.

Remark III.3. *Theorem III.1 and Example III.2 suggest the exponential generating function $\widetilde{\psi}_X$ is more amenable than the ordinary generating function ψ_X. We state both results for future reference, but recommend the first formulation.*

IV. De Bruijn identity

Motivated by [1], we give a de Bruijn identity with respect to beamsplitter addition $\boxplus_\eta$. The key result is the following discrete analogue of the heat equation, analogous to [7, Corollary 4.2] in the Poisson case:

Theorem IV.1. *For a given random variable X consider $Z_\eta := X \boxplus_\eta Y$, where Y is geometric. Writing $\lambda(\eta) = \lambda_{Z_\eta} = \eta\lambda_X + (1-\eta)\lambda_Y$ we obtain*

$$\frac{\partial}{\partial\eta} p_{Z_\eta}[n] := \Delta\left(\frac{n}{\eta}\left(p_{Z_\eta}[n-1]\lambda_Y - p_{Z_\eta}[n](1+\lambda_Y)\right)\right), \tag{15}$$

where for any function u, we write $\Delta(u[n]) := u[n+1] - u[n]$.

Proof: See Section IX. ∎

Definition IV.2. *For a random variable X with p.m.f. p_X, define two new p.m.f.s supported on $\mathbb{Z}_+$ by*

$$p_X^+[n] = \frac{(n+1)p_X[n+1]}{\lambda_X} \quad \text{and} \quad p_X^-[n] = \frac{(n+1)p_X[n]}{1+\lambda_X}. \tag{16}$$

This allows us to deduce the following de Bruijn identity, which is a specialization to number-diagonal states of the more general de Bruijn identity proved by König and Smith [13]:

Theorem IV.3. *Given $Z_\eta = X \boxplus_\eta Y$, where Y is geometric with mean λ_Y, we can write G_η for a geometric with mean $\lambda(\eta) = \eta\lambda_X + (1-\eta)\lambda_Y$. Then*

$$\frac{\partial}{\partial\eta} D(Z_\eta \| G_\eta) = \frac{\lambda_Y(1+\lambda(\eta))}{\eta} D(p_{Z_\eta}^- \| p_{Z_\eta}^+) + \frac{(1+\lambda_Y)\lambda(\eta)}{\eta} D(p_{Z_\eta}^+ \| p_{Z_\eta}^-),$$

where $p_{Z_\eta}^+$ and $p_{Z_\eta}^-$ are defined in terms of (16).

If X is itself geometric then so is Z_η, meaning that $p_{Z_\eta}^+ = p_{Z_\eta}^-$ (a negative binomial p.m.f.) and the two relative entropy terms on the RHS of Theorem IV.3 vanish as expected.

We focus on the case where $\lambda_X = \lambda_Y$, where the RHS of Theorem IV.3 becomes a symmetrized relative entropy. If G is geometric with $\mathbb{E}G = \lambda_X$ then direct calculation gives that

$$D(X\|G) = H(G) - H(X). \tag{17}$$

This allows us to deduce the following log-Sobolev type inequality which may be of independent interest:

Corollary IV.4. *For any random variable X, if G is geometric with mean $\mathbb{E}G = \lambda_X$ then:*

$$D(X\|G) \leq \lambda_X(1+\lambda_X)\left(D(p_X^- \| p_X^+) + D(p_X^+ \| p_X^-)\right). \tag{18}$$

Proof: We consider $Z_\eta = X \boxplus_\eta Y$, where Y is geometric with $\lambda_Y = \lambda_X$, and apply [12, Theorem 5], which tells us that $H(Z_\eta) \geq \eta H(X) + (1-\eta)H(Y)$. Combining this with (17), we can write

$$D(Z_\eta\|G) \leq \eta D(X\|G) + (1-\eta)D(Y\|G) = \eta D(X\|G),$$

or rearranging that (since $\eta \leq 1$)

$$\frac{D(Z_\eta\|G) - D(X\|G)}{\eta - 1} \geq D(X\|G).$$

If $\eta \to 1$, the LHS becomes the derivative $\frac{\partial}{\partial\eta} D(Z_\eta\|G)|_{\eta=1}$, and we deduce the result using Theorem IV.3. ∎

In the language of [5], Theorem IV.1 suggests that we can introduce a score function ρ_X, defined as:

Definition IV.5. *For a random variable X with p.m.f. p_X and mean λ_X, define a score function*

$$\rho_X[n] := \frac{n p_X[n-1]\lambda_X}{p_X[n](1+\lambda_X)} - n, \tag{19}$$

where we define $\rho_X[0] = 0$ to ensure $\sum_{n=0}^\infty p_X[n]\rho_X[n] = 0$. We define two Fisher-type quantities in terms of it:

$$J^+(X) := \sum_{n=1}^\infty \frac{p_X[n]}{n}\rho_X[n]^2, \tag{20}$$

$$J^-(X) := \sum_{n=1}^\infty \frac{p_X[n]}{n+\rho_X[n]}\rho_X[n]^2. \tag{21}$$

Note that ρ_X vanishes if and only if X is geometric, so that $J^+(X)$ and $J^-(X)$ are ≥ 0, with equality if and only if X is geometric. Further, if we choose Y to be geometric with mean $\mathbb{E}Y = \lambda_X$ then (15) becomes

$$\frac{\partial}{\partial\eta} p_{Z_\eta}[n] := \frac{1+\lambda_X}{\eta}\Delta\left(p_{Z_\eta}[n]\rho_{Z_\eta}[n]\right).$$

where as before, we write $\Delta(u[n]) := u[n+1] - u[n]$. Secondly, linearising the logarithm in Corollary IV.4 implies a quadratic version of this result, in the spirit of [15]:

Corollary IV.6. *For any random variable X, if G is geometric with mean $\mathbb{E}G = \lambda_X$ then:*

$$D(X\|G) \leq (1+\lambda_X)\left(J^+(X) + J^-(X)\right).$$

V. Log-sum inequality

We state one further result, which controls how the relative entropy behaves under the action of the map $\mathcal{T}$.

Theorem V.1. *Given two random variables X and Y, with $\mathbf{X}_c = \mathcal{T}(X)$ and $\mathbf{Y}_c = \mathcal{T}(Y)$ then:*

$$D(\mathbf{X}_c\|\mathbf{Y}_c) \leq D(X\|Y).$$

Proof: Writing $\phi_n(\mathbf{r}) := e^{-|\mathbf{r}|^2}|\mathbf{r}|^{2n}/(n!\pi)$ then, using (1) and the log-sum inequality [16, Theorem 2.7.1], for any $\mathbf{r}$:

$$\begin{aligned} p_{\mathbf{X}_c}(\mathbf{r}) \log\left(\frac{p_{\mathbf{X}_c}(\mathbf{r})}{p_{\mathbf{Y}_c}(\mathbf{r})}\right) &= \left(\sum_{n=0}^\infty p_X[n]\phi_n(\mathbf{r})\right)\log\left(\frac{\sum_{n=0}^\infty p_X[n]\phi_n(\mathbf{r})}{\sum_{n=0}^\infty p_Y[n]\phi_n(\mathbf{r})}\right) \\ &\leq \sum_{n=0}^\infty p_X[n]\phi_n(\mathbf{r})\log\left(\frac{p_X[n]}{p_Y[n]}\right). \end{aligned}$$

Integrating over $\mathbf{r}$ we deduce the result since $\int \phi_n(\mathbf{r})d\mathbf{r} = 1$ for each n. ■

VI. Conclusions

We have introduced a new purely classical representation for the beamsplitter addition operation $\boxplus_\eta$, with respect to the exponential generating function. We have deduced a heat equation, and recovered a purely classical proof of a special case of the de Bruijn identity of König and Smith [13].

In future work, we hope to use this formalism to prove discrete entropy results based around the geometric family, analogous to the classical results proved for the continuous entropy based around the Gaussian family. In particular, we hope that our results can give insights into a proof of the conjectured discrete EPI under beamsplitter addition [12]—a special case of the Entropy Photon Number Inequality [9], [10] —as well as give insights into convergence to the geometric and the conjectured monotonic increase in entropy under repeated beamsplitter addition [12], analogous to the classical Central Limit Theorem convergence to Gaussians and 'law of thin numbers' [6] convergence to Poissons.

VII. Proof of Transform relation, Theorem II.4

Proof of Lemma II.2:

1) We reverse the order of summation in (5) to obtain

$$\begin{aligned} \widetilde{\psi}_X(t) &= \sum_{m=0}^{\infty} \frac{t^m}{m!} \sum_{n=m}^{\infty} \binom{n}{m} p_X[n] \\ &= \sum_{n=0}^{\infty} p_X[n] \sum_{m=0}^{n} \binom{n}{m} \frac{t^m}{m!} \\ &= \sum_{n=0}^{\infty} p_X[n] \mathcal{L}_n(-t), \end{aligned} \tag{22}$$

since $\mathcal{L}_n(-t) = \sum_{m=0}^n \binom{n}{m} \frac{t^m}{m!}$ (see [17, Eq. (5.1.6)]).

2) This result follows on integrating $\exp(-s)\mathcal{L}_m(s)$ times both sides of (22) with $s = -t$, using the orthogonality relation for Laguerre polynomials [17, Eq. (5.1.1)],

$$\int_0^\infty \exp(-s)\mathcal{L}_m(s)\mathcal{L}_n(s)ds = \delta_{mn}.$$

■

Recall the standard result that the Laplace transform L relates exponential and ordinary generating functions:

Lemma VII.1. *Given a sequence $\mathbf{a} = (a_n)_{n=0,1,\ldots}$, if we write $\psi_{\mathbf{a}}(t) = \sum_{n=0}^\infty a_n t^n$ and $\widetilde{\psi}_{\mathbf{a}}(t) = \sum_{n=0}^\infty a_n t^n/n!$, then*

$$\left[L\widetilde{\psi}_{\mathbf{a}}\right](u) = \frac{1}{u}\psi_{\mathbf{a}}\left(\frac{1}{u}\right). \tag{23}$$

Proof: This follows since

$$\begin{aligned} \left[L\widetilde{\psi}_{\mathbf{a}}\right](u) &= \sum_{n=0}^{\infty} \frac{a_n}{n!} \left(\int_0^\infty \exp(-su)s^n ds \right) \\ &= \sum_{n=0}^{\infty} \frac{a_n}{n!} \left(\frac{n!}{u^{n+1}} \right) = \frac{1}{u}\psi_{\mathbf{a}}\left(\frac{1}{u}\right). \end{aligned} \tag{24}$$

■

Proof of Theorem II.4: We can express (7) in terms of the Bessel function J_0 (see [17, Eq. (1.71.1)]), which we substitute to obtain (25) below. We obtain:

$$\begin{aligned} \widetilde{\phi}_{\mathbf{X}_c}(-t) &= \int p_{\mathbf{X}_c}(\mathbf{r}) \sum_{m=0}^{\infty} \frac{|\mathbf{r}|^{2m}(-t)^m}{m!^2} d\mathbf{r} \\ &= \int p_{\mathbf{X}_c}(\mathbf{r}) J_0(2|\mathbf{r}|\sqrt{t}) d\mathbf{r} && (25) \\ &= \sum_{n=0}^{\infty} p_X[n] \int \frac{e^{-|\mathbf{r}|^2}|\mathbf{r}|^{2n}}{n!\pi} J_0(2|\mathbf{r}|\sqrt{t}) d\mathbf{r} && (26) \\ &= 2\sum_{n=0}^{\infty} p_X[n] \int_0^\infty \frac{e^{-r^2} r^{2n+1}}{n!} J_0(2r\sqrt{t}) dr && (27) \\ &= \sum_{n=0}^{\infty} p_X[n] \int_0^\infty \frac{e^{-u}u^n}{n!} J_0(2\sqrt{ut}) du && (28) \\ &= \sum_{n=0}^{\infty} p_X[n] \mathcal{L}_n(t)\exp(-t), && (29) \end{aligned}$$

where (26) follows by substituting (1), (27) follows by moving from Cartesian coordinates $d\mathbf{r}$ to polar $rdrd\theta$, (28) uses $u = r^2$ and (29) follows by [17, Theorem 5.4.1]. The result follows by Lemma II.2.

Consider taking Laplace transforms of both sides of (8). Using Lemma VII.1 the Laplace transform of the LHS is

$$\left[L\widetilde{\psi}_X\right](u) = \frac{1}{u}\psi_X\left(\frac{1}{u}\right), \tag{30}$$

Again by Lemma VII.1, since the Laplace transform of $f(t)\exp(-t)$ is the Laplace transform of f shifted by 1, the Laplace transform of the RHS is

$$\left[L\widetilde{\phi}_{\mathbf{X}_c}\right](u+1) = \frac{1}{u+1}\phi_{\mathbf{X}_c}\left(\frac{1}{1+u}\right). \tag{31}$$

Equating (30) and (31), the result follows taking $u = 1/t$. ■

VIII. Proof of convolution relation, Theorem III.1

We first state a result that shows how the moments of circularly symmetric random variables on $\mathbb{C}$ behave on convolution:

Lemma VIII.1. *Given independent circularly symmetric $\mathbf{X}_c$ and $\mathbf{Y}_c$ and writing $\mathbf{Z}_c := \sqrt{\eta}\mathbf{X}_c + \sqrt{1-\eta}\mathbf{Y}_c$, we can write*

$$\widetilde{\phi}_{\mathbf{Z}_c}(t) = \widetilde{\phi}_{\mathbf{X}_c}(\eta t)\,\widetilde{\phi}_{\mathbf{Y}_c}((1-\eta)t). \tag{32}$$

Proof: Consider independent $\mathbf{U}_c$ and $\mathbf{V}_c$ and write $\mathbf{W}_c = \mathbf{U}_c + \mathbf{V}_c$. Then [18, Eq. (3)] gives

$$\mathbb{E}|\mathbf{W}_c|^{2m} = \sum_{n=0}^{m} \binom{m}{n}^2 \mathbb{E}|\mathbf{U}_c|^{2n}\mathbb{E}|\mathbf{V}_c|^{2m-2n}.$$

Multiplying $t^m/(m!)^2$, and summing, we obtain that

$$\widetilde{\phi}_{\mathbf{W}_c}(t) = \widetilde{\phi}_{\mathbf{U}_c}(t)\widetilde{\phi}_{\mathbf{V}_c}(t)$$

and the result follows by rescaling. ■

Putting all this together we obtain:

Proof of Theorem III.1:

1. This result follows directly on combining (8) and (32).

2. Relabelling $t = 1/(s-1)$ in (9) and using Lemma VII.1, for any random variable U we obtain:

$$\left[L\widetilde{\phi}_{\mathbf{U}_c}\right](s) = \frac{1}{s}\phi_{\mathbf{U}_c}\left(\frac{1}{s}\right) = \frac{1}{s-1}\psi_U\left(\frac{1}{s-1}\right). \quad (33)$$

Taking $U = Z$ in (33), and using Lemma VIII.1, we know that $\widetilde{\phi}_{\mathbf{Z}_c} = M^{(\eta)}_{\mathbf{X}_c} \times M^{(1-\eta)}_{\mathbf{Y}_c}$ where $M^{(\eta)}_{\mathbf{X}_c}(t) = \widetilde{\phi}_{\mathbf{X}_c}(\eta t)$. We can use the fact that if $F = [Lf]$ then $[Lf(at)](s) = \frac{1}{a}F\left(\frac{s}{a}\right)$ to deduce using (33) that

$$\left[LM^{(\eta)}_{\mathbf{X}_c}\right](t) = \frac{1}{\eta}\left[L\widetilde{\phi}_{\mathbf{X}_c}\right]\left(\frac{t}{\eta}\right) = \frac{1}{\eta}\frac{1}{s-1}\psi_X\left(\frac{1}{s-1}\right)\bigg|_{s=t/\eta}$$

and (12) follows. A similar argument based on the fact that $M^{(1-\eta)}_{\mathbf{Y}_c}(t) = \widetilde{\phi}_{\mathbf{Y}_c}((1-\eta)t)$ allows us to deduce (13). ■

IX. Proof of de Bruijn identity, Theorem IV.3

We first prove the heat equation Theorem IV.1:

Proof of Theorem IV.1: By Lemma II.2.1) we can write

$$h(\eta;t) := \widetilde{\psi}_{Z_\eta}(t) = \sum_{n=0}^{\infty} p_{Z_\eta}[n]\mathcal{L}_n(-t). \quad (34)$$

Using (10) we also write $h(\eta;t) = \widetilde{\psi}_X(\eta t)\exp((1-\eta)\lambda_Y t)$ and observe that this satisfies

$$\frac{\partial}{\partial \eta}h(\eta;t) = \frac{t}{\eta}\frac{\partial}{\partial t}h(\eta;t) - \frac{\lambda_Y t}{\eta}h(\eta;t). \quad (35)$$

Hence, by differentiating (34) and using (35), we obtain

$$\begin{aligned}
&\sum_{n=0}^{\infty}\frac{\partial}{\partial\eta}p_{Z_\eta}[n]\mathcal{L}_n(-t) \\
&= \frac{t}{\eta}\frac{\partial}{\partial t}h(\eta;t) - \frac{\lambda_Y t}{\eta}h(\eta;t) \\
&= \frac{-t}{\eta}\sum_{n=0}^{\infty}p_{Z_\eta}[n]\Big(\mathcal{L}'_n(-t) + \lambda_Y\mathcal{L}_n(-t)\Big) \\
&= -\frac{\lambda_Y}{\eta}\sum_{n=0}^{\infty}p_{Z_\eta}[n]\Big((n+1)\mathcal{L}_{n+1}(-t) - (n+1)\mathcal{L}_n(-t)\Big) \\
&\quad + \frac{1+\lambda_Y}{\eta}\sum_{n=0}^{\infty}p_{Z_\eta}[n]\Big(n\mathcal{L}_n(-t) - n\mathcal{L}_{n-1}(-t)\Big) \quad (36) \\
&= \sum_{n=0}^{\infty}\Delta\left(\frac{n}{\eta}\left(p_{Z_\eta}[n-1]\lambda_Y - p_{Z_\eta}[n](1+\lambda_Y)\right)\right)\mathcal{L}_n(-t)
\end{aligned}$$

here (36) follows using the fact that $z\mathcal{L}'_n(z) = n\mathcal{L}_n(z) - n\mathcal{L}_{n-1}(z)$ (see [17, Eq. (5.1.14)]) and using the three-term relation for Laguerre polynomials, $-z\mathcal{L}_n(z) = (n+1)\mathcal{L}_{n+1}(z) - (2n+1)\mathcal{L}_n(z) + n\mathcal{L}_{n-1}(z)$ (see [17, Eq. (5.1.10)]). Comparing coefficients of $\mathcal{L}_n(-t)$ we conclude the result holds. ■

Proof of Theorem IV.3: Using (17), and writing p_η for p_{Z_η}, we can express $D(Z_\eta\|G_\eta)$ as

$$\sum_{n=0}^{\infty}p_\eta[n]\log p_\eta[n] - \lambda(\eta)\log\lambda(\eta) + (1+\lambda(\eta))\log(1+\lambda(\eta)).$$

For any function $u[n]$, the $\sum_{n=0}^{\infty}\Delta(u[n])\log p_\eta[n] = \sum_{n=0}^{\infty}u[n+1]\log(p_\eta[n]/p_\eta[n+1])$ if $u[0]=0$, so (assuming we can exchange the sum and derivative) Theorem IV.1 gives

$$\begin{aligned}
&\frac{\partial}{\partial\eta}D(Z_\eta\|G_\eta) \\
&= \sum_{n=0}^{\infty}\left(\frac{\partial}{\partial\eta}p_\eta[n]\right)\log p_\eta[n] - \lambda'(\eta)\log\left(\frac{\lambda(\eta)}{1+\lambda(\eta)}\right) \\
&= \sum_{n=0}^{\infty}\frac{n+1}{\eta}\left(p_\eta[n]\lambda_Y - p_\eta[n+1](1+\lambda_Y)\right) \\
&\qquad \times\log\left(\frac{p_\eta[n]/(1+\lambda(\eta))}{p_\eta[n+1]/\lambda(\eta)}\right) \quad (37)
\end{aligned}$$

where (37) follows using (14). Inserting factors of $(n+1)$ in the top and bottom of the fraction we deduce the result. ■

Acknowledgements: The authors thank the organizers of the 2016 Beyond i.i.d. in Information Theory workshop, where discussions leading up to this paper took seed.

References

[1] A. J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Inform. Control*, vol. 2, pp. 101–112, 1959.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[3] A. R. Barron, "Entropy and the Central Limit Theorem," *Ann. Probab.*, vol. 14, no. 1, pp. 336–342, 1986.

[4] M. Madiman and A. R. Barron, "Generalized entropy power inequalities and monotonicity properties of information," *IEEE Trans. Inform. Theory*, vol. 53, no. 7, 2317 – 2329, 2007.

[5] I. Kontoyiannis, P. Harremoës, and O. T. Johnson, "Entropy and the law of small numbers," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 466–472, 2005.

[6] P. Harremoës, O. T. Johnson and I. Kontoyiannis, "Thinning, Entropy and the Law of Thin Numbers," *IEEE Trans. Inform. Theory*, vol. 56, no.9, pp. 4228–4244, 2010.

[7] O. T. Johnson, "Log-concavity and the maximum entropy property of the Poisson distribution," *Stoch. Proc. Appl.*, vol. 117, no. 6, pp. 791–802, 2007.

[8] O. T. Johnson and Y. Yu, "Monotonicity, thinning and discrete versions of the Entropy Power Inequality," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5387–5395, 2010.

[9] S. Guha, J. H. Shapiro, and B. I. Erkmen, "Capacity of the bosonic wiretap channel and the entropy photon-number inequality," in *2008 International Symposium on Information Theory*, IEEE, pp. 91–95, 2008.

[10] S. Guha, B. I. Erkmen, and J. H. Shapiro, "The entropy photon-number inequality and its consequences," in *Information Theory and Applications Workshop, 2008.* pp. 128–130, 2008.

[11] H. Qi, M. M. Wilde, and S. Guha, "Thermal states minimize the output entropy of single-mode phase-insensitive Gaussian channels with an input entropy constraint," see arxiv:`1607.05262`, 2016.

[12] S. Guha, J. H. Shapiro, and R. G. Sanchez, "Thinning, photonic beamsplitting, and a general discrete entropy power inequality," see `arxiv:1605.07853`, 2016.

[13] R. König and G. Smith, "The entropy power inequality for quantum systems," *IEEE Trans. Inform. Theory*, vol. 60, no. 3, pp. 1536–1548, 2014.

[14] B. Rimoldi, *Principles of Digital Communication*. Cambridge University Press, 2016.

[15] S. G. Bobkov and M. Ledoux, "On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures," *J. Funct. Anal.*, vol. 156, no. 2, pp. 347–365, 1998.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley, 1991.

[17] G. Szegő, *Orthogonal Polynomials*, revised ed. New York: American Mathematical Society, 1958.

[18] J. Goldman, "Moments of the sum of circularly symmetric random variables" *IEEE Trans. Inform. Theory*, vol. 18, no. 2, pp. 297–298, 1972.